

Методы сегментации спутниковых изображений для оконтуривания речных выносов с помощью нейронных сетей

А. Н. Якушева, Н. А. Князев, М. В. Врублевский

Институт космических исследований РАН, Москва, 117997, Россия

E-mail: annayaakusheva99@mail.ru

Настоящая работа посвящена разработке и исследованию методов автоматической сегментации спутниковых изображений для оконтуривания речных выносов с использованием свёрточных нейронных сетей. Исследование проводилось на основе анализа данных мультиспектральных сенсоров MSI (*англ.* Multispectral Instrument) спутников Sentinel-2 для четырёх ключевых районов: устьевых зон рек Сулак и Терек в Каспийском море, р. Мзымты в Чёрном море и Калининградского канала в Балтийском море. В работе был создан собственный набор данных, включающий 587 фрагментов спутниковых изображений с экспертной разметкой речных плутов. Проведено комплексное сравнение 10 различных архитектур нейронных сетей, включая комбинации энкодеров (VGG16, ResNet50, Xception) с декодерами (U-Net, DeepLabv3+, FCN), а также модель YOLO 11-Seg. Для повышения качества сегментации разработан программный комплекс предварительной обработки данных, включающий модули загрузки, кадрирования и разметки спутниковых изображений. Экспериментальные исследования показали, что архитектура YOLO 11-Seg продемонстрировала наилучшую производительность с коэффициентом Дайса 0,74. Для дальнейшего улучшения результатов были реализованы два метода ансамблирования: взвешенное усреднение масок и стекинг с использованием Random Forest в качестве метамодели. Стекинг-подход показал наилучшие результаты, достигнув коэффициента Дайса 0,78. Разработанная методология демонстрирует высокую эффективность автоматического детектирования речных выносов в различных гидрологических условиях и может быть интегрирована в системы оперативного спутникового мониторинга морской среды.

Ключевые слова: спутниковый мониторинг, MSI Sentinel-2, речные выносы, нейросети, сегментация, Сулак, Терек, Мзымта, Калининградский залив

Одобрена к печати: 18.06.2025

DOI: 10.21046/2070-7401-2025-22-3-81-94

Введение

Исследование речных выносов по спутниковым данным представляет собой важную научную задачу, имеющую как фундаментальное, так и прикладное значение. Речные выносы, содержащие взвешенные вещества, питательные элементы и загрязнители, играют ключевую роль в формировании экосистем прибрежных вод, влияя на биологическую продуктивность и седиментационные процессы. Традиционные методы изучения речных выносов, основанные на натурных измерениях, обладают существенными ограничениями, связанными с их локальным характером, высокой трудоёмкостью и невозможностью обеспечить систематическое покрытие обширных акваторий.

Развитие спутниковых технологий дистанционного зондирования Земли открыло новые возможности для решения этой задачи. Мультиспектральные сенсоры современных космических аппаратов, таких как Sentinel-2, обеспечивают регулярный мониторинг водной поверхности с высоким пространственным и спектральным разрешением. Это позволяет исследовать пространственно-временную динамику речных выносов в ранее недоступных масштабах.

Значительный вклад в развитие методов спутникового мониторинга речных выносов внесли исследования, проведённые в Институте космических исследований РАН. В работах (Лаврова и др., 2022; Назирова и др., 2019, 2023) были описаны эффективные методики визуального анализа и ручного дешифрирования речных плутов по данным спутниковых наблюдений. Эти подходы, основанные на экспертной интерпретации изображений, доказали свою

надёжность, однако требуют значительных временных затрат и не поддаются полной автоматизации, что ограничивает их применение для оперативного мониторинга больших акваторий.

Современные технологии искусственного интеллекта, в частности методы глубокого обучения для анализа изображений, предлагают принципиально новые возможности для автоматизации процесса детектирования речных выносов. Пороговая сегментация спутниковых снимков с использованием свёрточных нейронных сетей позволяет существенно ускорить обработку данных при сохранении высокой точности выделения целевых объектов. Такой подход особенно перспективен для задач оперативного мониторинга, где критически важны скорость обработки и воспроизводимость результатов.

В настоящей работе мы исследуем возможность применения методов нейросетевой сегментации для автоматизированного детектирования речных выносов по данным спутника Sentinel-2. Основное внимание уделяется Каспийскому морю, в частности зонам влияния рек Сулак и Терек, Чёрному морю, где расположена устьевая зона р. Мзымты, и Калининградскому каналу — месту выноса внутренних вод из Калининградского зал. в Балтийское море. Эти реки характеризуются высокой мутностью вод и формируют протяжённые плюмы, динамика которых остаётся недостаточно изученной. Выбор Sentinel-2 в качестве источника данных обусловлен уникальной комбинацией высокого пространственного разрешения (до 10 м) и частоты съёмки (5 дней), а также наличием спектральных каналов, оптимальных для детектирования взвешенных веществ. Особое внимание в исследовании уделяется сравнению результатов автоматической обработки с данными экспертного анализа, что позволяет объективно оценить эффективность предлагаемого подхода.

Методология исследования включает несколько ключевых этапов: подготовку обучающей выборки на основе экспертных разметок, обучение модели сегментации, валидацию результатов и их сравнение с традиционными методами анализа. Особое внимание уделяется вопросам обеспечения точности и воспроизводимости результатов, что имеет принципиальное значение для практического применения разработанных методов.

Районы исследования

Каспийское море представляет собой уникальный природный объект для изучения процессов формирования и распространения речных выносов. Особый научный интерес вызывает его северо-западная часть, куда поступают воды крупных рек Кавказа, формируя обширные зоны смешения речных и морских вод. Этот район отличается исключительно высокой динамичностью гидрологических процессов, обусловленной сочетанием нескольких ключевых факторов: интенсивным речным стоком, сложной батиметрией шельфовой зоны, специфическими ветровыми режимами и выраженной сезонной изменчивостью гидрометеорологических условий (Kostianoy et al., 2019).

Гидрологический режим Северо-Западного Каспия в значительной степени определяется стоком двух крупных рек — Терека и Сулака, которые совместно дают около 25 % всего речного стока в море. Река Терек протяжённостью в 623 км, берущая начало на склонах Главного Кавказского хр., отличается преобладанием ледникового питания (до 70 % стока) и характерным летним половодьем (рис. 1, см. с. 83). В период с июня по август она выносит в море огромное количество взвешенных наносов, формируя протяжённые мутные шлейфы. Особенностью р. Терек является её крайне изменчивая дельта, где русловые процессы приводят к постоянной перестройке системы протоков, что существенно влияет на пространственное распределение выносов.

Река Сулак протяжённостью в 144 км образуется при слиянии Аварского Койсу и Андийского Койсу (рис. 2, см. с. 83). В весенний паводковый период (апрель — май) концентрация взвешенных веществ достигает наибольших показателей, что связано с интенсивным размытием глинистых пород в среднем течении реки. Формируемый р. Сулак мутный плюм хорошо различим на спутниковых снимках и в благоприятных гидрометеороусловиях может проследиться на расстоянии 30–40 км от устья.



Рис. 1. Приустьевая зона р. Терек. Фрагмент изображения MSI (англ. Multispectral Instrument) Sentinel-2B от 12 июня 2023 г.



Рис. 2. Приустьевая зона р. Сулак. Фрагмент изображения MSI Sentinel-2A от 17 июня 2023 г.

Сезонная динамика речных выносов в Каспийском районе отличается исключительной контрастностью. Весенне-летний период (апрель – июль) даёт до 70 % годового объёма выносов, что связано с таянием снегов и ледников в бассейнах рек. Именно в это время формируются наиболее протяжённые и насыщенные взвешенными веществами речные плюмы, оказывающие максимальное влияние на экосистему моря. Преобладающие в этот период северо-западные ветры способствуют вдольбереговому движению выносов в юго-восточном направлении, создавая сложные пространственные структуры распределения мутности. Летом интенсивность выносов несколько снижается, но остаётся значительной благодаря ливневым паводкам, тогда как осенне-зимний период характеризуется минимальными объёмами стока наносов.

В работе также исследуется Черноморское побережье Кавказа, а именно устьевая зона р. Мзымты, характеризующаяся интенсивным выносом взвеси в прибрежной зоне (рис. 3, см. с. 84). До 80 % объёма от годового стока наносов реки поступает в море в течение весенне-летнего периода (апрель – июль), что связано с таянием сезонных снегов и ледников в высокогорной части бассейна, а также ливневыми паводками. Особенностью выноса р. Мзымты является формирование протяжённого мутного шлейфа, распространяющегося на 15–20 км от устья, с выраженной пространственно-временной изменчивостью, обусловленной сочетанием речного стока, морских течений и ветрового воздействия. Концентрация взвешенных веществ в зоне выноса может достигать 50–100 NTU (англ. Nephelometric Turbidity Unit, нефелометрическая единица мутности) в период паводков, что оказывает существенное влияние на оптические свойства прибрежных вод.



Рис. 3. Приустьевая зона р. Мзымты. Фрагмент изображения MSI Sentinel-2B от 11 июня 2023 г.

Ещё один объект исследования — зона выноса Калининградского зал. Калининградский (Вислинский) зал., соединяющийся с Балтийским морем через Балтийский пролив, является важной целью для анализа процессов речного выноса в условиях ограниченного водообмена с открытым морем (рис. 4).

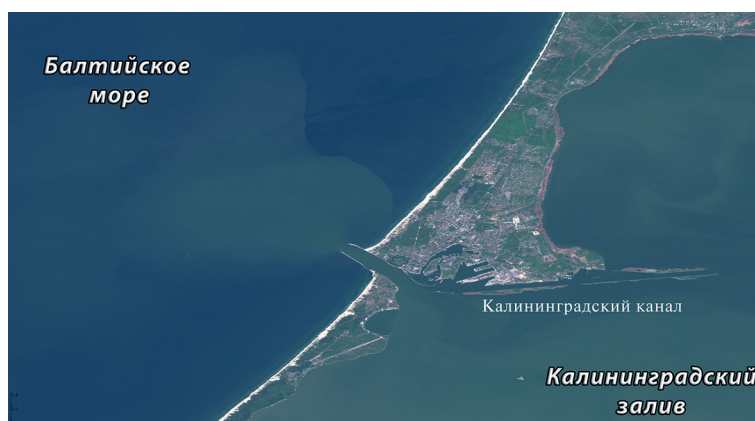


Рис. 4. Калининградский морской канал. Фрагмент изображения MSI Sentinel-2A от 7 мая 2023 г.

Этот объект представляет научный интерес как пример антропогенно-трансформированного выноса в условиях ограниченного водообмена. Мониторинг динамики выноса залива важен для понимания процессов эвтрофикации прибрежных вод, оценки влияния хозяйственной деятельности (судоходство, промышленные и коммунальные стоки) на качество вод, а также разработки мер по устойчивому управлению прибрежными экосистемами. Сравнительный анализ выносов различных типов (горные реки Кавказа и равнинный сток Балтийского региона) позволяет выявить общие закономерности формирования речных плюмов и разработать универсальные подходы к их мониторингу с использованием методов дистанционного зондирования.

Используемые спутниковые данные

В работе использованы данные оптических сенсоров мультиспектральных датчиков MSI космических аппаратов Sentinel-2A/2B высокого пространственного разрешения (10 м), распространяемые в рамках программы Copernicus (англ. Copernicus Open Access Hub) (<https://>

scihub.copernicus.eu/). Основные характеристики выбранных для дальнейшего анализа каналов спутниковых изображений приведены в *табл. 1*. Выбор конкретных каналов обоснован двумя причинами. Во-первых, применяемые нейросетевые методы требуют трёхканальных изображений, поэтому мы вынуждены ограничиться лишь тремя каналами (Chollet, 2017; He et al., 2016; Redmon et al., 2016; Simonyan, Zisserman, 2014). Во-вторых, речные выносы наиболее ярко проявляются в оптическом диапазоне, соответствующем 2, 3 и 4-му каналу MSI, что подтверждается многолетним опытом и выработанной методикой ряда специалистов (Лаврова и др., 2011). Всего в работе было использовано 587 фрагментов спутниковых изображений MSI.

Таблица 1. Основные характеристики используемых каналов MSI

Каналы MSI	Спектральный диапазон, мкм	Пространственное разрешение, м
2	0,46–0,52	10
3	0,54–0,58	
4	0,65–0,68	

Подготовка данных

Для проведения сегментации выносов на оптических данных спутниковых снимков нейросетевыми методами было необходимо получить и предварительно подготовить данные. С этой целью были сформулированы следующие задачи. На основе имеющейся базы данных (БД) наблюдений речных выносов, обнаруженных и оконтуренных в системе спутникового мониторинга STS (*англ.* See the Sea) (Лупян и др., 2012) вручную, проводить опрос каталога данных Copernicus на наличие снимков в области интереса (границах выноса из БД) за соответствующие даты наблюдения. Осуществлять загрузку найденных данных. Вести отчётность о полноте загруженных файлов; об обнаружении или необнаружении данных по критериям поиска, соответствующих наблюдению выносов из БД STS; об ошибках, которые могут возникнуть на любом этапе работы по получению данных. Осуществлять выбор необходимых каналов из набора скачанных данных. Кадрировать спутниковый снимок, сохраняя только территорию интереса с целью экономии занимаемой памяти и ускорения нейросетевого обучения. Размечать набор данных, формируя соответствие между пикселями спутникового изображения и границами выносов из БД STS.

Для решения поставленных задач был подготовлен программный комплекс (ПК) получения и предварительной подготовки данных (разметки, выбора каналов, кадрирования данных снимков оптических сенсоров). Общая архитектура ПК показана на *рис. 5*.

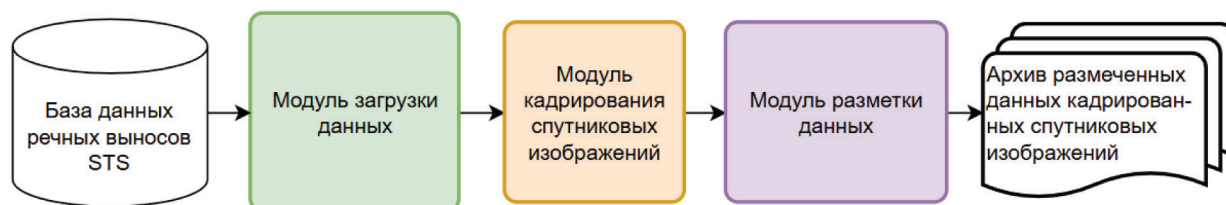


Рис. 5. Общая архитектура программного комплекса получения и предварительной подготовки данных

В состав ПК входит ряд программных модулей:

- Модуль загрузки данных, предназначенный для загрузки данных и выполняющий следующие функции: опрос каталога для определения наличия спутниковых данных, относящихся к границам выносов из БД STS за соответствующие даты наблюдения;

загрузка данных из каталога Copernicus; проверка целостности загруженных файлов; реализация повторной загрузки файлов в случае возникновения каких-либо ошибок при предыдущей попытке загрузки; ведение отчётов об ошибках и загрузке.

- Модуль кадрирования спутниковых изображений, предназначенный для экономии пространства памяти и сокращения времени работы нейросетевого метода сегментирования спутниковых изображений, выполняющего следующие функции: формирование границ для кадрирования спутникового снимка за счёт создания прямоугольного буфера (бокса) к геометрии, заданной границами речного выноса из БД STS; объединение фрагментов сцены спутниковой съёмки в одно изображение; кадрирование спутникового изображения с удалением исходных файлов.
- Модуль разметки данных, предназначенный для бинарной классификации (вынос/не вынос) изображений на основе границ речных выносов из БД STS, выполняющий функции: растеризация векторной геометрии границ речных выносов в виде бинарной маски (вынос/не вынос) спутниковых изображений; объединение в один массив данных полученной маски и необходимых для дальнейших действий каналов спутникового изображения; проведение операции сжатия данных; формирование архива размеченных данных кадрированных спутниковых изображений.

Реализация описанного ПК позволила подготовить архив данных, необходимый для дальнейшей работы с данными с помощью нейросетевых методов, а также сократить объём памяти, требуемый для хранения архива файлов, с приблизительно 1,2 Тб до менее 15 Гб, что существенно сказалось на времени, необходимом для нейросетевого обучения. Оптимизация формата хранения данных обеспечила 7–8-кратное ускорение обучения для всех тестируемых архитектур сегментации. Работа ПК осуществлялась в среде Google Colab с использованием языка программирования Python.

Разработанный ПК за счёт модульной структуры можно быстро адаптировать под работу с другими каталогами спутниковых данных, каналами оптических изображений, форматами хранения данных на регулярной сетке, сенсорами в качестве источника данных (например, радиолокационными изображениями).

Основные компоненты нейронных сетей

Энкодер нейросети для сегментации кодирует входные данные, преобразуя их в компактное представление с меньшим количеством параметров (LeCun et al., 2015). В компьютерном зрении эту роль выполняют свёрточные слои. Свёртка (*англ.* convolution) выделяет пространственные признаки через весовые фильтры, а пулинг (*англ.* pooling) агрегирует локальные значения, уменьшая разрешение (Long et al., 2015). Комбинация этих операций формирует иерархию признаков от простых градиентов до семантических шаблонов (Zeiler, Fergus, 2014).

В настоящей работе использовались энкодеры VGG16, ResNet50, Xception. Энкодер VGG16 (Simonyan, Zisserman, 2014) увеличивает глубину представлений при уменьшении пространственного разрешения, эффективно захватывая признаки различной сложности; ResNet50 (He et al., 2016) вводит остаточные блоки (*англ.* Residual blocks) и пропуски соединений (*англ.* skip-connections), решая проблему затухания градиентов и улучшая восстановление пространственной структуры. Архитектура Xception (Chollet, 2017) использует глубинно-сепарабельные свёртки (*англ.* Depthwise separable convolutions), разделяя пространственную и межканальную обработку для эффективного выделения многоуровневых признаков.

Декодер выполняет обратную энкодеру задачу — восстановление исходных данных из скрытого вектора через нелинейные преобразования, увеличивающие размерность (LeCun et al., 2015). В компьютерном зрении декодеры реализуются через транспонированные свёртки (*англ.* deconvolution), выполняющие операцию обратную стандартной свёртке, или развёртку (*англ.* upsampling) с интерполяцией (Long et al., 2015).

В данном исследовании применялись декодеры U-Net, DeepLabv3+ и FCN. Декодер U-Net (Ronneberger et al., 2015) построен симметрично энкодеру: транспонированные свёртки и конкатенации карт признаков через пропуски соединений для восстановления границ объектов; DeepLabv3+ (Chen et al., 2018) использует модуль ASPP (*англ.* Atrous Spatial Pyramid Pooling), обрабатывающий изображение в разных масштабах через изменение коэффициентов дилатации (шаг свёртки); FCN (Long et al., 2015) применяет каскад транспонированных свёрток с пропусками соединений для повышения разрешения.

Также в настоящей работе использовалась модель **YOLO 11-Seg** — эволюция семейства YOLO (*англ.* You Only Look Once). В данной модели сохранён базовый принцип: свёрточный энкодер разбивает изображение на сетку ячеек для регрессии ограничивающих прямоугольников (*англ.* bounding boxes) и определения вероятностей классов (Bochkovskiy et al., 2020). Модель YOLO 11-Seg использует двуступенчатый декодер с каскадом FPA (*англ.* Feature Pyramid Attention), формирующий маски различного разрешения с последующей агрегацией через пропуски соединения (Kirillov et al., 2019).

Подготовка данных для нейросетей

Для решения задач с ограниченными данными применялось трансферное обучение — метод повторного использования предобученных моделей (Pan, Yang, 2010). Использовались веса VGG16, ResNet50, Xception и YOLO 11-Seg, обученные на миллионах изображений. При применении предобученных моделей необходимо соблюдать оригинальную нормализацию данных для стабильного обучения, поскольку без неё значения пикселей $[0, 255]$ создают слишком большие градиенты.

Для VGG16, ResNet50, Xception и YOLO 11-Seg используются схожие принципы нормализации. Основной подход заключается в приведении значений пикселей к диапазону $[0, 1]$ путём деления на 255, после чего применяется стандартизация по каналам RGB с использованием средних значений и стандартных отклонений набора данных ImageNet (Deng et al., 2009), на котором обучались модели.

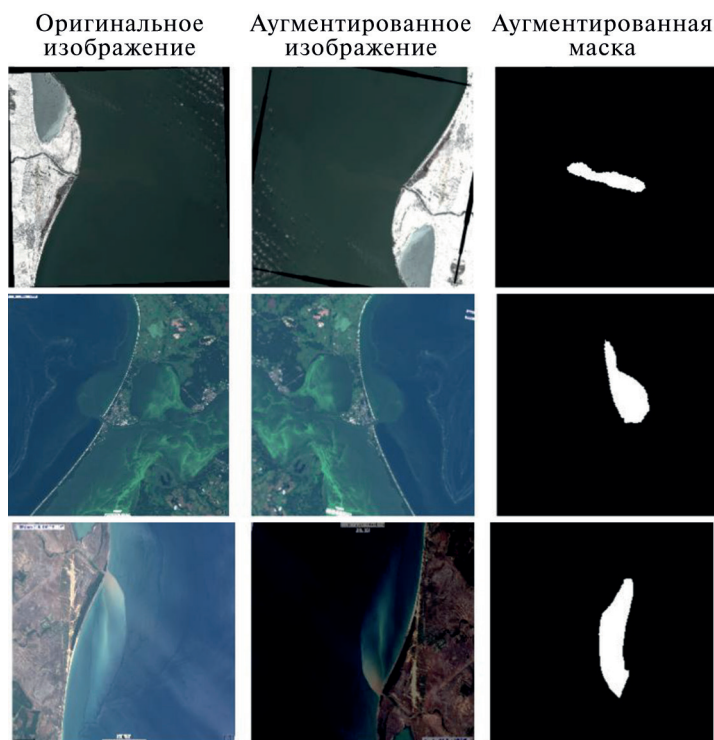


Рис. 6. Аугментация данных

Следующим важным этапом подготовки данных является аугментация — процедура искусственного увеличения набора данных через преобразования изображений. В задачах сегментации аугментация применяется синхронно к изображению и маске. В настоящей работе использовались геометрические и фотометрические преобразования.

Геометрические преобразования применялись синхронно к изображению и маске: горизонтальное и вертикальное отражение с вероятностью $p = 0,5$, случайное вращение на угол $\theta \in [-15^\circ, 15^\circ]$. Для заполнения областей после поворота использовались билинейная интерполяция для изображений и интерполяция ближайшего соседа для масок (сохранение дискретных значений класса).

В обучающем наборе содержится 410 изображений речных выносов, в валидационном — 126, а в тестовом — 51. Важно отметить, что аугментация данных (рис. 6, см. с. 87), в целях предотвращения переобучения, применялась только к обучающему и валидационному, но не тестовому набору, чтобы окончательная оценка качества осуществлялась исключительно на данных фактических наблюдений.

Результаты обучения

В рамках данного исследования технология трансферного обучения применялась двухэтапно: энкодеры инициализировались с предобученными весами с заморозкой начальных слоёв для сохранения универсальных паттернов, затем проводилась тонкая настройка всей архитектуры для адаптации к сегментации спутниковых снимков. Проводилась комплексная оценка YOLO 11-Seg и гибридных моделей энкодеров (VGG16, ResNet50, Xception) с декодерами (DeepLabv3+, FCN, U-Net).

Схема экспериментов охватывала все возможные комбинации энкодер-декодер ($3 \times 3 = 9$ конфигураций) и отдельное применение YOLO 11-Seg. При этом обучение всех моделей проводилось на едином наборе данных с использованием идентичных гиперпараметров для обеспечения корректности сравнительного анализа.

Оценка сегментации использует метрики на основе истинно положительных TP (*англ.* True Positive), FP ложно положительных (*англ.* False Positive), ложно отрицательных FN (*англ.* False Negative), истинно отрицательных TN (*англ.* True Negative) исходов. Метрика точности (*англ.* Accuracy) — доля правильных пикселей. Но данная метрика некорректна при дисбалансе классов. Другой метрикой является точность (*англ.* Precision) — доля истинно положительных среди предсказанных положительных. Метрика полноты (*англ.* Recall) показывает выявление истинных положительных исходов. Коэффициент Дайса (*англ.* Dice coefficient) — главная метрика, представляющая гармоническое среднее точности и полноты. Эта метрика чувствительна к небольшим объектам и дисбалансу классов. Значения выше 0,7 приемлемы для сегментации.

Также в ходе данного исследования были протестированы различные функции потерь для оптимизации обучения моделей сегментации: коэффициент Дайса, Focal Loss для трудных примеров и комбинированные функции. Экспериментально установлено, что бинарная кросс-энтропия обеспечивает наилучшее качество работы и стабильную сходимость. Для оптимизации весов использовался алгоритм Adam (Kingma, Ba, 2014), сочетающий адаптивную скорость обучения и эффективную обработку градиентов.

Результаты сравнения архитектур энкодер-декодер (табл. 2) (полужирное начертание означает лучшие результаты в сочетании энкодер-декодер) для данной задачи сегментации демонстрируют значительные различия в качестве работы между различными комбинациями архитектур. Анализ метрик показывает, что выбор архитектуры энкодера и декодера существенно влияет на качество сегментации.

Модель YOLO 11-Seg показала наилучшие результаты среди всех протестированных архитектур, достигнув коэффициента Дайса 0,74. Энкодер ResNet50 продемонстрировал наилучший результат ($\text{Dice} = 0,69$) в паре с декодером DeepLabv3+, что подтверждает эффективность комбинирования остаточных связей с многомасштабной обработкой ASPP-модуля. Самое хорошее качество работы энкодера Xception наблюдается с деко-

дером FCN (Dice = 0,60). Лучшие показатели энкодера VGG16 достигнуты с декодером U-Net (Dice = 0,66), что можно объяснить совместимостью симметричной архитектуры U-Net с относительно простой структурой VGG16. Сравнение декодеров показывает, что DeepLabv3+ обеспечивает самое высокое качество работы благодаря эффективному использованию многомасштабных признаков. Наилучший баланс данных метрик достигается в архитектуре YOLO 11-Seg, что подтверждает её превосходство в задачах сегментации.

Таблица 2. Результаты обучения

		Декодеры			Метрики
		U-net	DeepLabv3+	FCN	
Энкодеры	ResNet50	0,64 0,98/0,82/0,60	0,69 0,98/0,76/0,72	0,67 0,98/0,78/0,66	Dice Accuracy/Precision/Recall
	Xception	0,47 0,97/0,57/0,45	0,52 0,97/0,52/0,63	0,60 0,97/0,68/0,61	Dice Accuracy/Precision/Recall
	VGG16	0,66 0,98/0,74/0,67	0,56 0,97/0,61/0,66	0,55 0,98/0,64/0,56	Dice Accuracy/Precision/Recall
	YOLO 11-Seg	0,74 0,99/0,72/0,84			Dice Accuracy/Precision/Recall

Результаты сравнительного анализа (рис. 7, 8, см. с. 90) наилучших сочетаний архитектур энкодер-декодер для задач сегментации демонстрируют неоднозначную картину качества работы. На представленных изображениях видно, что каждая модель имеет свои сильные и слабые стороны в зависимости от характеристик объектов сегментации и условий съёмки. DeepLabv3+ с энкодером ResNet50 показывает стабильные результаты на всех тестовых изображениях, обеспечивая чёткое выделение основных контуров. Архитектура эффективно справляется с задачей сегментации речных выносов.

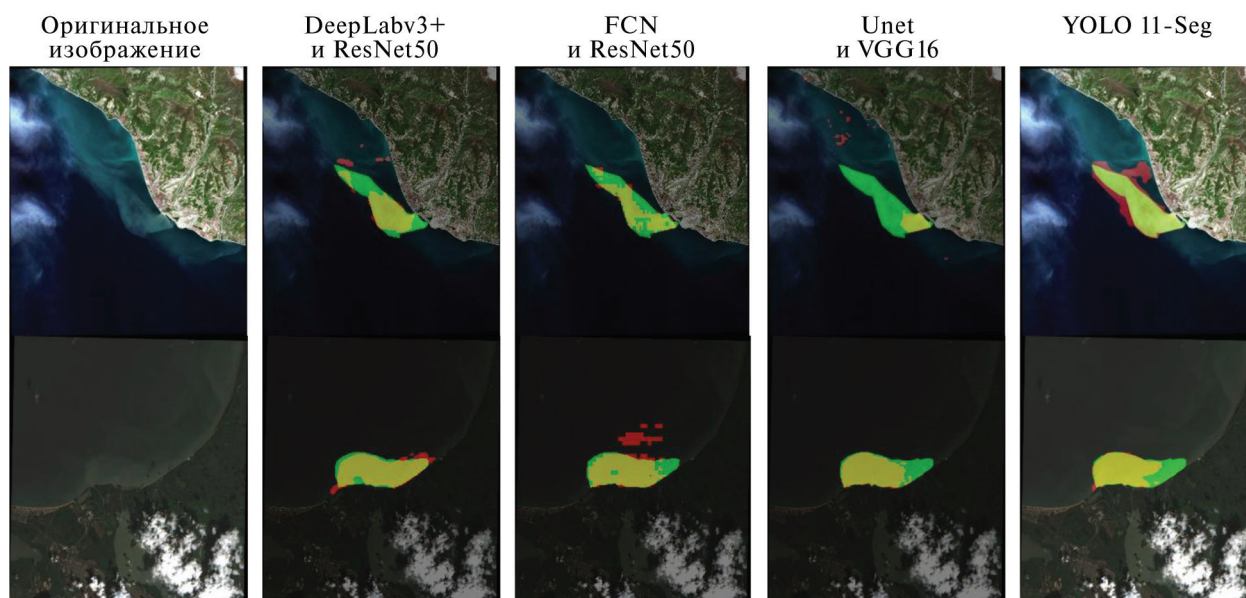


Рис. 7. Результаты сегментации выносов. Зелёная область — маска плюма, размеченная экспертом; красная область — маска, отмеченная нейросетью; жёлтая область — пересечение масок

Модель FCN с энкодером ResNet50 преимущественно выделяет наиболее уверенно классифицируемые экспертами области. Декодер U-Net с энкодером VGG16 захватывает небольшие области. Архитектура YOLO 11-Seg, несмотря на наивысшие показатели точности в ряду

метрик, демонстрирует неоднозначные результаты при визуальном анализе. В некоторых случаях наблюдается включение областей, которые не соответствуют размеченным экспертами границам речных плумов. Это подчёркивает важность комплексной оценки моделей, поскольку высокие численные метрики не всегда коррелируют с качеством сегментации в сложных случаях.

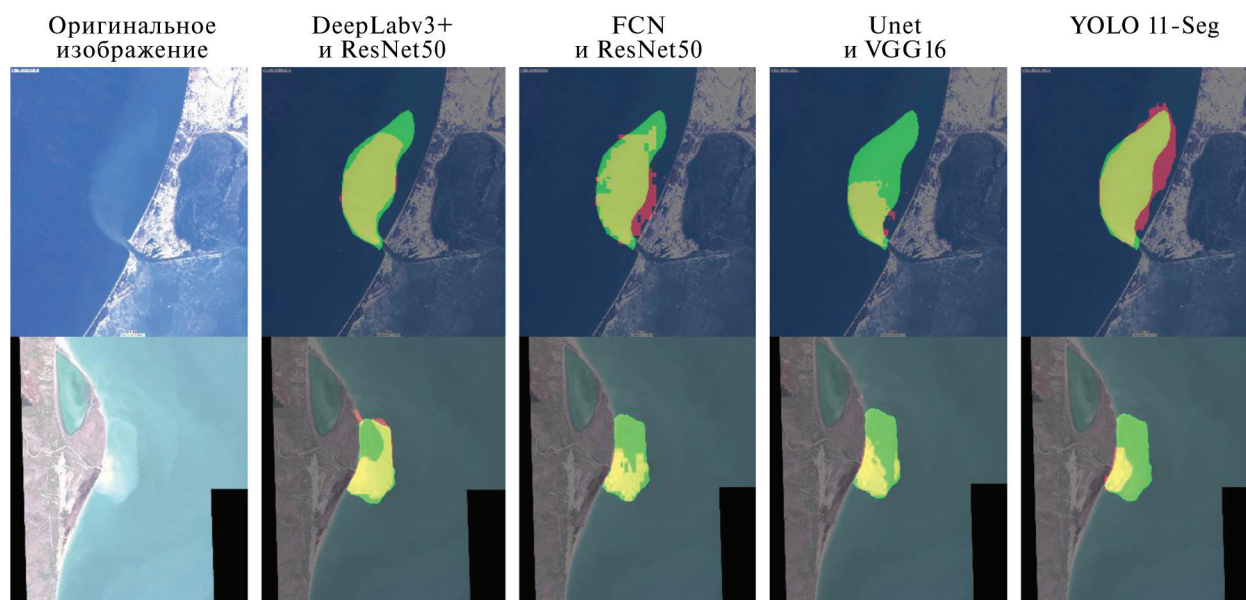


Рис. 8. Результаты сегментации выносов. Зелёная область — маска плюма, размеченная экспертом; красная область — маска, отмеченная нейросетью; жёлтая область — пересечение масок

Оптимальный выбор архитектуры нейронной сети для задачи сегментации определяется комплексом факторов, включающих специфические требования конкретного изображения. Основываясь на детальном визуальном анализе результатов сегментации, было принято решение объединить предсказания четырёх базовых моделей (U-Net, DeepLabv3+, FCN и YOLO 11-Seg) для повышения качества сегментации плумов. Для этого были реализованы два различных подхода к ансамблированию.

Первый метод — взвешенное усреднение масок. Данный подход заключается в подборе оптимальных весовых коэффициентов для каждой модели и последующем вычислении взвешенного среднего предсказанных масок. Для каждого пикселя итоговая маска вычисляется как комбинация предсказаний всех базовых моделей с учётом их индивидуальной важности, определённой эмпирически на основе производительности каждой архитектуры. Коэффициент Дайса данного метода составил 0,77.

Второй метод — стекинг (*англ.* stacking), который представляет собой технику ансамблирования, где предсказания базовых моделей (метапризнаки) используются в качестве входных признаков для обучения метамоделей. В отличие от простого усреднения, стекинг позволяет метаалгоритму изучить сложные нелинейные зависимости между предсказаниями различных моделей. В качестве метамоделей был выбран алгоритм Random Forest благодаря его способности эффективно работать с разнородными признаками и устойчивости к переобучению. Обучение стекинг-модели проводилось на масках, восстановленных наилучшими архитектурами на валидационном наборе данных, что обеспечивает корректность процедуры и предотвращает утечку данных. Такой подход позволяет метамоделю изучить паттерны ошибок и сильных сторон каждой базовой модели, формируя более точные и надёжные итоговые предсказания сегментации. Коэффициент Дайса данного метода составил 0,78.

Результаты стекинг-подхода (рис. 9, см. с. 91) демонстрируют высокую эффективность автоматического обнаружения и сегментации речных плумов для всех районов иссле-

дования. На представленных изображениях показано сравнение экспертной разметки (зелёные маски) с предсказаниями модели Random Forest (красные маски) и пересечения этих масок (жёлтые маски).

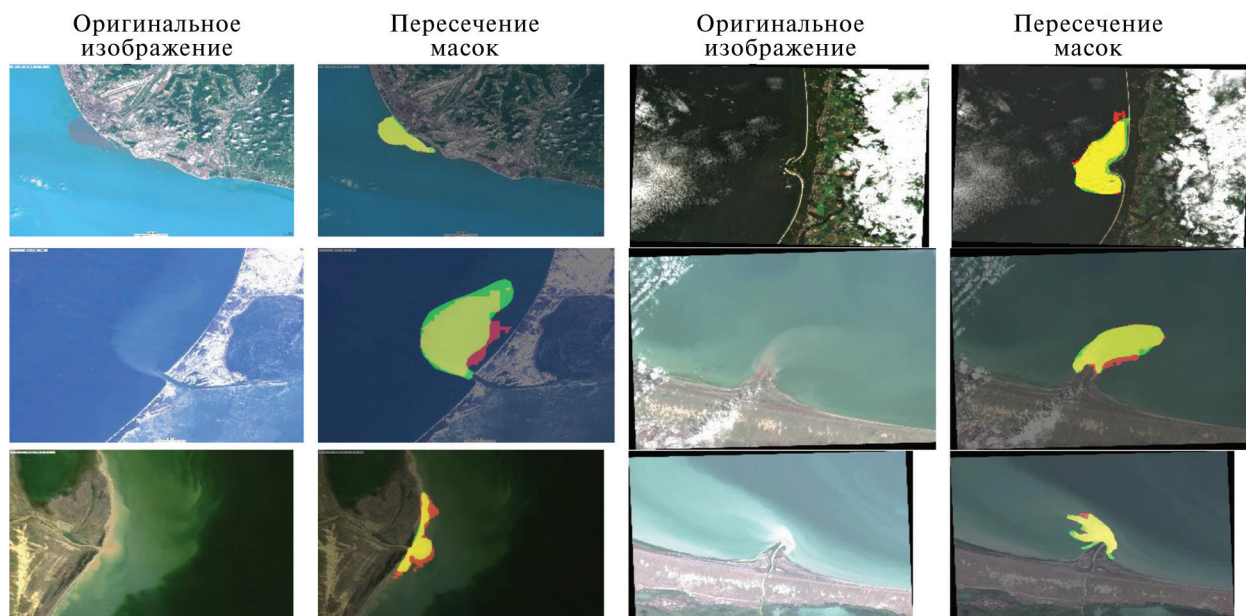


Рис. 9. Результаты стекинга

Анализ результатов демонстрирует, что стекинг-ансамбль успешно выявляет различные типы плумов независимо от их морфологических характеристик. Особенно заметна способность модели корректно сегментировать сложные случаи. Пересечение масок указывает на высокое соответствие между экспертной разметкой выносов и восстановленными масками. Способность данной системы демонстрирует наилучший подход к различным условиям освещения и типам загрязнения, что подтверждает эффективность комбинирования множественных энкодеров-декодеров для данной задачи. Точность локализации границ загрязнений позволяет использовать данный подход для оперативного реагирования и долгосрочного мониторинга качества прибрежных вод.

Выводы

Проведённое исследование продемонстрировало эффективность применения глубоких нейронных сетей для задач сегментации речных выносов на спутниковых изображениях в разных регионах исследования. В ходе работы была создана и обработана собственная база данных изображений спутниковых наблюдений.

Экспериментальная часть включала обучение и сравнение 10 различных архитектур нейронных сетей с разнообразными энкодерами и декодерами. Наилучший результат показала модель на основе архитектуры YOLO 11-Seg, достигшая коэффициента Дайса 0,74 на тестовой выборке, что свидетельствует о высокой точности сегментации речных плумов.

Для дальнейшего повышения качества алгоритма был разработан специализированный подход, основанный на обученных нейронных сетях, который позволил улучшить показатели сегментации до значения коэффициента Дайса 0,78. Это улучшение на 5,4 % демонстрирует потенциал применения комбинированных методов для специфических задач экологического мониторинга.

Полученные результаты позволяют произвести внедрение разработанного алгоритма в систему спутникового мониторинга See the Sea. Дальнейшая работа предполагает масшта-

бирование системы и адаптацию к различным условиям наблюдения. Также предстоит работа по улучшению результатов с помощью специализированных методов сегментации для объектов с размытыми и нечёткими границами, поскольку даже экспертам крайне сложно точно определить границы плюма. Полученные в настоящем исследовании результаты показывают высокое качество автоматической сегментации, сопоставимое с возможностями человеческой экспертизы. Проведённый анализ демонстрирует значительный потенциал методов машинного обучения для создания автоматизированных систем мониторинга речных выносов. Разработанный подход может быть использован как для фундаментальных исследований процессов речных выносов, так и для решения прикладных задач экологического мониторинга прибрежных зон. Полученные результаты особенно актуальны в контексте развития методов оперативного спутникового мониторинга морской среды.

Исследование выполнено за счёт гранта Российского научного фонда № 24-17-00182 «Развитие методов дистанционной диагностики распространения речных вод в прибрежной зоне морей» (<https://rscf.ru/project/24-17-00182/>) в Институте космических исследований РАН.

Литература

1. Лаврова О. Ю., Костяной А. Г., Лебедев С. А., Митягина М. И., Гинзбург А. И., Шеремет Н. А. Комплексный спутниковый мониторинг морей России. М.: ИКИ РАН, 2011. 480 с.
2. Лаврова О. Ю., Назирова К. Р., Алферьева Я. О. и др. Сопоставление параметров плюмов рек Сулак и Терек на основе спутниковых данных и измерений *in situ* // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 5. С. 264–283. DOI: 10.21046/2070-7401-2022-19-5-264-283.
3. Лупян Е. А., Матвеев А. А., Уваров И. А. и др. Спутниковый сервис See the Sea — инструмент для изучения процессов и явлений на поверхности океана // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 9. № 2. С. 251–261.
4. Назирова К. Р., Лаврова О. Ю., Краюшкин Е. В. и др. Особенности выявления параметров речного плюма контактными и дистанционными методами // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 2. С. 227–243. DOI: 10.21046/2070-7401-2019-16-2-227-243.
5. Назирова К. Р., Лаврова О. Ю., Алферьева Я. О., Князев Н. А. Пространственно-временная изменчивость плюмов рек Терек и Сулак по спутниковым данным и синхронным натурным измерениям // Современные проблемы дистанционного зондирования Земли из космоса. 2023. Т. 20. № 5. С. 285–303. DOI: 10.21046/2070-7401-2023-20-5-285-303.
6. Bochkovskiy A., Wang C.-Y., Liao H.-Y. M. YOLOv4: Optimal speed and accuracy of object detection // <https://arxiv.org/>. 2020. 17 p. DOI: 10.48550/arXiv.2004.10934.
7. Chen L.-C., Zhu Y., Papandreou G. et al. Encoder-decoder with atrous separable convolution for semantic image segmentation // Proc. 15th European Conf. Computer Vision — ECCV 2018. Pt. 7. Lecture Notes in Computer Science. V. 11211. Cham: Springer, 2018. P. 833–851. DOI: 10.1007/978-3-030-01234-2_49.
8. Chollet F. Xception: Deep learning with depthwise separable convolutions // Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2017). 2017. P. 1800–1807. DOI: 10.1109/CVPR.2017.195.
9. Deng J., Dong W., Socher R. et al. ImageNet: A large-scale hierarchical image database // Proc. 2009 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2009). 2009. P. 248–255. DOI: 10.1109/CVPR.2009.5206848.
10. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2016). 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.
11. Kingma D. P., Ba J. Adam: A method for stochastic optimization // <https://arxiv.org/>. 2014. 15 p. DOI: 10.48550/arXiv.1412.6980.
12. Kirillov A., Girshick R., He K., Dollár P. Panoptic Feature Pyramid Networks // Proc. 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR 2019). 2019. P. 6392–6401. DOI: 10.1109/CVPR.2019.00656.
13. Kostianoy A. G., Ginzburg A. I., Lavrova O. Yu. et al. Comprehensive satellite monitoring of Caspian Sea conditions // Remote Sensing of the Asian Seas / eds. Barale V., Gade M. Cham: Springer Intern. Publ., 2019. P. 505–521. DOI: 10.1007/978-3-319-94067-0_28.
14. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. V. 521. P. 436–444. DOI: 10.1038/nature14539.

14. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. 2015. V. 521. P. 436–444. DOI: 10.1038/nature14539.
15. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation // Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2015). 2015. P. 3431–3440. DOI: 10.1109/CVPR.2015.7298965.
16. Pan S. J., Yang Q. A survey on transfer learning // IEEE Trans. Knowledge and Data Engineering. 2010. V. 22. No. 10. P. 1345–1359. DOI: 10.1109/TKDE.2009.191.
17. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, real-time object detection // Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2016). 2016. P. 779–788. DOI: 10.1109/CVPR.2016.91.
18. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Pt. 3. Lecture Notes in Computer Science. V. 9351. Cham: Springer, 2015. P. 234–241. DOI: 10.1007/978-3-319-24574-4_28.
19. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // <https://arxiv.org/>. 2014. 14 p. DOI: 10.48550/arXiv.1409.1556.
20. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks // Proc. 13th European Conf. Computer Vision — ECCV 2014. Pt. 1. Lecture Notes in Computer Science. V. 8689. Cham: Springer, 2014. P. 818–833. DOI: 10.1007/978-3-319-10590-1_53.

Satellite image segmentation methods for river plume delineation using neural networks

A. N. Yakusheva, N. A. Knyazev, M. V. Vrublevsky

Space Research Institute RAS, Moscow 117997, Russia

E-mail: annayaakusheva99@mail.ru

The work is devoted to the development and study of methods of automatic segmentation of satellite images using convolutional neural networks with the aim of river plume delineation. The study was based on analysis of data from Sentinel-2 MSI multispectral sensors for four main regions: the mouth zones of the Sulak and Terek rivers in the Caspian Sea, the Mzymta River in the Black Sea and the Kaliningrad Canal in the Baltic Sea. A dataset comprising 587 fragments of satellite images with expert marking of river plumes was created. A comprehensive comparison of 10 different neural network architectures was conducted, including combinations of encoders (VGG16, ResNet50, Xception) with decoders (U-Net, DeepLabv3+, FCN), as well as the YOLO 11-Seg model. To improve the quality of segmentation, a software package for data pre-processing was developed, including modules for loading, cropping and labeling satellite images. Experimental studies showed that the YOLO 11-Seg architecture demonstrated the best performance with a Dice coefficient of 0.74. To further improve the results, two ensembling methods were implemented: weighted mask averaging and stacking using Random Forest as a metamodel. The stacking approach showed the best results, achieving a Dice coefficient of 0.78. The developed methodology demonstrates high efficiency of automatic detection of river plumes in various hydrological conditions and can be integrated into operational satellite monitoring systems of the marine environment.

Keywords: satellite monitoring, Sentinel-2 MSI, river plume, neural networks, segmentation, Sulak, Terek, Mzymta, Kaliningrad Bay

Accepted: 18.06.2025

DOI: 10.21046/2070-7401-2025-22-3-81-94

References

1. Lavrova O. Yu., Kostianoy A. G., Lebedev S. A., Mityagina M. I., Ginzburg A. I., Sheremet N. A., *Kompleksnyi sputnikovyi monitoring morei Rossii* (Complex satellite monitoring of Russian seas), Moscow: IKI RAS, 2011, 480 p. (in Russian).

2. Lavrova O. Yu., Nazirova K. R., Alferyeva Ya. O. et al., Comparison of plume parameters of the Sulak and Terek rivers based on satellite data and *in situ* measurements, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2022, V. 19, No. 5, pp. 264–283 (in Russian), DOI: 10.21046/2070-7401-2022-19-5-264-283.
3. Loupian E. A., Matveev A. A., Uvarov I. A. et al., The Satellite Service See the Sea — a tool for the study of oceanic phenomena and processes, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, V. 9, No. 2, pp. 251–261 (in Russian).
4. Nazirova K. R., Lavrova O. Yu., Krayushkin E. V. et al., Features of river plume parameter determination by *in situ* and remote sensing methods, *Sovremennye problem distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, V. 16, No. 2, pp. 227–243 (in Russian), DOI: 10.21046/2070-7401-2019-16-2-227-243.
5. Nazirova K. R., Lavrova O. Yu., Alferyeva Ya. O., Knyazev N. A., Spatiotemporal plume variability of Terek and Sulak rivers from satellite data and concurrent *in situ* measurements, *Sovremennye problem distantsionnogo zondirovaniya Zemli iz kosmosa*, 2023, V. 20, No. 5, pp. 285–303 (in Russian), DOI: 10.21046/2070-7401-2023-20-5-285-303.
6. Bochkovskiy A., Wang C.-Y., Liao H.-Y. M., YOLOv4: Optimal speed and accuracy of object detection, <https://arxiv.org/>, 2020, 17 p., DOI: 10.48550/arXiv.2004.10934.
7. Chen L.-C., Zhu Y., Papandreou G. et al., Encoder-decoder with atrous separable convolution for semantic image segmentation, In: *Proc. 15th European Conf. Computer Vision — ECCV 2018*, Pt. 7, Lecture Notes in Computer Science, V. 11211, Cham: Springer, 2018, pp. 833–851, DOI: 10.1007/978-3-030-01234-2_49.
8. Chollet F., Xception: Deep learning with depthwise separable convolutions, *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2017)*, 2017, pp. 1800–1807, DOI: 10.1109/CVPR.2017.195.
9. Deng J., Dong W., Socher R. et al., ImageNet: A large-scale hierarchical image database, *Proc. 2009 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2009)*, 2009, pp. 248–255, DOI: 10.1109/CVPR.2009.5206848.
10. He K., Zhang X., Ren S., Sun J., Deep residual learning for image recognition, *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2016)*, 2016, pp. 770–778, DOI: 10.1109/CVPR.2016.90.
11. Kingma D. P., Ba J., Adam: A method for stochastic optimization, <https://arxiv.org/>, 2014, 15 p., DOI: 10.48550/arXiv.1412.6980.
12. Kirillov A., Girshick R., He K., Dollár P., Panoptic Feature Pyramid Networks, *Proc. 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR 2019)*, 2019, pp. 6392–6401, DOI: 10.1109/CVPR.2019.00656.
13. Kostianoy A. G., Ginzburg A. I., Lavrova O. Yu. et al., Comprehensive satellite monitoring of Caspian Sea conditions, In: *Remote Sensing of the Asian Seas*, Barale V., Gade M. (eds.), Cham: Springer Intern. Publ., 2019, pp. 505–521, DOI: 10.1007/978-3-319-94067-0_28.
14. LeCun Y., Bengio Y., Hinton G., Deep learning, *Nature*, 2015, V. 521, pp. 436–444, DOI: 10.1038/nature14539.
15. Long J., Shelhamer E., Darrell T., Fully convolutional networks for semantic segmentation, *Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2015)*, 2015, pp. 3431–3440, DOI: 10.1109/CVPR.2015.7298965.
16. Pan S. J., Yang Q., A survey on transfer learning, *IEEE Trans. Knowledge and Data Engineering*, 2010, V. 22, No. 10, pp. 1345–1359, DOI: 10.1109/TKDE.2009.191.
17. Redmon J., Divvala S., Girshick R., Farhadi A., You Only Look Once: Unified, real-time object detection, *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR 2016)*, 2016, pp. 779–788, DOI: 10.1109/CVPR.2016.91.
18. Ronneberger O., Fischer P., Brox T., U-Net: Convolutional networks for biomedical image segmentation, In: *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015*, Pt. 3, Lecture Notes in Computer Science, V. 9351, Cham: Springer, 2015, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28.
19. Simonyan K., Zisserman A., Very deep convolutional networks for large-scale image recognition, <https://arxiv.org/>, 2014, 14 p., DOI: 10.48550/arXiv.1409.1556.
20. Zeiler M. D., Fergus R., Visualizing and understanding convolutional networks, In: *Proc. 13th European Conf. Computer Vision — ECCV 2014*, Pt. 1, Lecture Notes in Computer Science, V. 8689, Cham: Springer, 2014, pp. 818–833, DOI: 10.1007/978-3-319-10590-1_53.